The Philosophy of the Law Idea and the Role of the Prescientific in Statistical Inference

Andrew M Hartley, PhD
Senior Biostatistician
UCB Pharma, Inc.
1421 Gate View Way
Marietta, GA 30062-2174

1. Introduction

> "[S]tatistics. . . (is) concerned with . . . methods and techniques for collecting, analyzing, and interpreting quantitative data in such a way that the reliability of conclusions based on the data may be evaluated objectively in terms of probability statements."[1]
> "It is proper to combine prior opinion and knowledge with p-values to guess the truth."[2]

Statistical analyses are routinely applied to substantiate conclusions in a wide range of sciences. Yet, they often reach contradictory results, because they usually follow two conflicting ideals. No decisive defenses have been proposed for their methods.

The first statement above promotes what might be called an "objectivist" ideal. As I will explain, this ideal leads to the deriving of statistical outcomes (such as p-values) based exclusively on quantitative observations. These outcomes are interpreted rigidly, as probabilities of hypotheses or as decision functions. Such interpretations are considered objective, in the sense of being "sought in empirical facts and logical reasoning."[3]

The second statement epitomizes a relatively subjectivist ideal. It advocates the generation of many of those same outcomes. However, it regards them merely as so-called "evidence" to be combined somehow with "prior opinion and knowledge." Scientific conclusions, from this perspective, take the form of personal, imprecise "guesses," rather than definitive or objective probabilities.

Clashes between these ideals and the conclusions they generate have sparked much research on statistical principles in recent years. Nonetheless, no clear direction for statistics has emerged.

The Philosophy of the Law Idea (PLI), on the other hand, has been promoted[4] as a biblically consistent account of experience. It deals with the roles in science of "prior opinion and knowledge," information I will call "prescientific." After reviewing the PLI's basic tenets, I suggest what it may indicate about the relation of statistical inference to the contexts in which statistics is applied. This provides clues concerning the objectivity and subjectivity we can expect in applied statistical inference.[5]

2. The Philosophy of the Law Idea

The PLI is rich in implications for the scope and methodology of all sciences. Its most important premise is that all of "reality is created by God whose will is the sovereign and redeeming law for reality."[6] It claims, on that basis, that every scientific pursuit, Christian and otherwise, is grounded on underlying religious presuppositions.[7]

2.1 Aspectual Irreducibility

In philosophy, it is generally agreed that the objects[8] of our pre-theoretical experience exhibit different aspects, or kinds of properties and laws. The PLI posits fifteen or so aspects:[9] quantitative, spatial, kinematic, physical, biological, sensory, logical, historical, linguistic, social, economic, aesthetic, jural, ethical and fiduciary.

According to the PLI, none of the aspects can be reduced to any other, in the senses either of being eliminated in favor of another or of being caused by it. We may not consider kinematic movement, for example, as only an "infinite series of discrete spatial magnitudes,"[10] or mathematics, which focuses on the quantitative aspect, as nothing but a logical process.[11]

Many Christians have regarded theories acceptable that causally reduce some aspects to others so long as the reducing aspects are said to depend in turn upon God. The PLI objects, however, claiming that all the aspects of creation depend on God intimately and directly, and that only Christ mediates God's sustaining power to creation (Col 1:15). Elevating any aspects above others is at odds with these teachings.

2.2 Aspectual Interconnectedness

Clouser explains that

> the aspects cannot be isolated from one another; their very intelligibility depends on their connectedness. Though they may be abstracted from the things which exhibit them, they cannot—even in thought—be isolated from one another. So even though the meanings of "quantitative," "physical," "sensory," "justitial," etc., are all importantly distinct and irreducible, they can only be understood in connection with, and by being compared to, one another. [12]

Therefore, we should not expect scientific theories to explain data completely logically or quantitatively. Each science can convey meaning only when it recognizes, explicates and refers to a multiplicity of aspects; "[n]o matter how hard a science may try to exclude all but its delimiting aspect, it cannot avoid dealing with the [properties] its data display in the other aspects."[13]

2.3 The PLI and Scientific-Prescientific

The PLI's distinction between the prescientific and scientific types of experience, and the relations it prescribes between them, are fundamental to its implications for statistics.

In prescientific life, we experience each thing as a concrete whole that displays simultaneously a plurality of aspects in mutual coherence.[14] The properties of the thing are certainly all present, but we do not take special note of any of them to the exclusion of others. "In daily life events, acts, things etc., are experienced as temporal totalities in which the different aspects are never experienced explicitly . . . but only implicitly."[15] Prescientific experience is not subjected to the "confrontation" that brings individual aspects into view.

In science, though, we single out (abstract) specific aspects.[16] We focus on a particular side of reality, heightening our ability to learn about particular properties and their relations within it.[17]

Some sciences are devoted to constructing explanations within a single aspect: physics - physical; language - linguistic; etc. Other sciences, though, such as archeology, deal with two or more aspects.[18] I propose that statistical inference is such a multi-aspectual science. For, I take it to be the quantitative assessment of the fiduciary credibility of the logical hypotheses of the sciences in which it is applied;[19] therefore, it must deal with the quantitative, fiduciary and logical aspects, as well as the aspects studied by those sciences.

The PLI implies that only prescientific experience can apprehend reality in the structure it gives itself, because only this attitude is attuned to that reality. Abstractive science can exist, then, only within the "non-dissected" structure of prescientific experience. It is not primary but secondary, a product of theoretical analysis. It pulls something essential away from the structure of reality. The prescientific differs from but is not inferior to science. It is valid and worthwhile, and provides the only proper foundation for science.

The PLI stands against two other, more prevalent attitudes about scientific-prescientific that I review next.

2.3.1 Scientism

"Scientism," according to the PLI, is an exaggeration of science's position in life. It is a "reversal of things" in which "[l]ife and experience, which always precede science, are second; science, which always must be second, is first."[20] Descartes' scientistic *more geometrico*, for instance,

> consists in having the scientific (geometrical) reason call all previous experience (i.e. experience previous to the application of the scientific method) into question until that scientific reason itself should discover some scientifically ascertainable absolute starting point for experience. From this scientifically fixed starting-point the reason . . . proceeds to build up . . . a new experience that would meet the

> scientific test. This latter experience, then . . . is the genuine experience. Here is the world and the word of truth. This is what we mean by scientism's reversal of the natural order. It *replaces* the experience of a life-time and the practical wisdom of the ages.[21]

Scientism claims that the few aspects abstracted by a particular science are sufficient for explaining multi-aspectual life. By assuming life can be reduced to these aspects, it violates the "irreducibility" principle.

The PLI criticizes not science as such, but rather any pretension that science independent of multi-aspectual prescientific experience can arrive at comprehensive truth.

2.3.2 Subjectivism

Some subjectivists agree with scientism that prescientific experiences "lack 'general validity' or 'objectivity'" and therefore need not be reckoned with.[22] They concur that abstraction is the proper starting-place for "the scientific method," and that science presents "an un-problematical datum."

These subjectivists ascribe self-sufficiency to science; still, their overall conception of science differs from the scientistic one. They claim the "chaotic . . . sensory-psychical impressions which form only the 'material' or 'stuff' of our experience"[23]. . . "are fed into the mind," and the mind in turn "orders them into an intelligible experience."[24] In other words, they locate the order of experience in the human intellect, not in any external world. They insist that humans exercise final, sovereign and subjective judgment respecting how to evaluate abstract scientific outcomes. This outlook prevents mathematical and logical rules from "transgressing into the normative sphere"[25] and interfering with human freedom.

The PLI opposes this subjectivism. It claims people (subjects) are not free to form whatever conclusions they like based on data (objects): "[I]n being known, the object places limitations on the subject who does the knowing,"[26] [27] including limitations on subjects' certainties. Science is a "confrontation" in which these limitations on subjects are discerned.

3. Nature of Applied Statistics

3.1 A Definition of Applied Statistics

To discuss the PLI and applied statistical inference with any precision, we must settle on a common understanding of applied statistics. I suggest defining applied statistics as

> the science of obtaining quantitative conclusions in the presence of uncertainty.[28]

Each "conclusion" in this sense can take one of two forms. First, it can be a statistical inference, meaning an assessment of the tenability (credibility) of a hypothesis.[29] [30] Alternatively, in a decision-making situation, a conclusion can be an assessment of the likely outcomes of a potential decision (allocation of resources). In this paper, I will focus on statistical inferences.

These inferences, though quantitative, concern phenomena with physical, biological, social, legal or other non-quantitative properties as well. The sciences in which statistical methods are applied define and study these properties. In other words, statistics is always applied in a context.[31]

3.2 Applied Statistics is Inductive

Researchers in the sciences reason inductively. Chemists, for instance, investigate general laws about molecules based on observations and on other accepted laws. It is not surprising, then, that statistics, when applied within the other sciences, must also proceed inductively.

Our definition of applied statistics above, along with many statistics texts, reflect the inductive character of this science. Huntsberger,[32] for example, defines statistical inference as "generalizing from the part to the whole. Given incomplete information (from the sample), [it] is making a statement about the larger group from which the sample was taken." *A fortiori*, Anderson[33] says more inclusively "statistics [itself] is one means of inferring generalities from specific observations, and it is objective" (echoing, again, the objectivity ideal).

4. An Example

We may illustrate the inductive intent of applied statistics, and the objectivity and subjectivity possible in its statements, by tracing two possible approaches to an application. They correspond to and illustrate two broad classes of statistical paradigms I will consider later.

A scientist wants to assess whether a particular object has an attribute A (hypothesis H_1) or not (hypothesis H_0). She submits the object to a test for detecting A. The test returns a positive result ($+T$).

She knows that, in this situation, $Pr(+T| H_1)$, the probability of $+T$ assuming H_1, is 0.80, and $Pr(-T| H_0)=0.95$. Thus, whatever the test result, neither H_1 nor H_0 is established unequivocally. The best she can aim for is some partial level of certainty in H_1.

4.1 Interpreting the Test Results: Inductive Approach

It is tempting to conclude,[34] based on $Pr(+T| H_1)=0.80$ and the result $+T$, that the scientist can be 80% certain of H_1. This conclusion would be objective (depend only on the observation $+T$), but is it justified?

In general, making statements about hypotheses such as H_1 based on observations such as $+T$ is scientific induction. One inductive statement relating directly to the scientist's question is $Pr(H_1|+T)$. Letting $\pi=Pr(H_1)$ be a pre-test probability[35] of H_1, one can show[36] $Pr(H_1|+T)=0.8\pi / (0.05+0.75\pi)$.

Two features of this inductive analysis are noteworthy. First, $Pr(H_1|+T)$ depends not only on $Pr(+T|\ H_1)$ and $Pr(-T|\ H_0)$, but on $\pi=Pr(H_1)$, which itself must be based on prescientific judgments such as the prevalence of A in the general population. $Pr(H_1|+T)$ is determined only insofar as π is determined. For instance, the $Pr(H_1|+T)$ associated with π=10% is 64%, and that associated with π=50% is 94%. Evidently, we cannot mathematically justify probabilistic statements about H_1 and H_0 without reference to the prescientific.

Second, $Pr(H_1|+T)$ prescribes (constrains) directly one's level of certainty that H_1 is true. Clinging nonetheless to a level of certainty about H_1 that is different from $Pr(H_1|+T)$ violates the accepted laws of probability.[37]

4.2 Interpreting the Test Results: Deductive Approach

Scientific deduction reasons from assumed hypotheses to observations. It proceeds in the direction opposite, so to speak, that of induction. It relates to the scientist's question indirectly at most. The statements $Pr(+T|\ H_1)=0.80$ and $Pr(-T|\ H_0)=0.95$ above are deductive, and no mathematical basis exists for reading them inductively. For instance, as we've seen, the inductive $Pr(H_1|+T)$ based on □ above can differ substantially from $Pr(+T|\ H_1)$.

Scientists do, however, speak about H_1 and H_0 referring only to deductive probabilities such as $Pr(+T|\ H_1)$ and $Pr(-T|\ H_0)$. Some, despite lack of mathematical justification, take $Pr(+T|\ H_1)$ and $Pr(H_1|+T)$ to be equal, so that $Pr(H_1|+T)=80\%$ in a "direct" manner.[38] Others more "indirectly" treat $+T$ as evidence about H_1 and H_0. One may combine informally this evidence with prescientific information. These two deductive kinds of attempts at induction illustrate the objectivistic and subjectivistic ideals mentioned at the outset of this paper.

5. Statistical Paradigms

The inductive and deductive methods above of interpreting $+T$ correspond to three general methods of interpreting data that I review next, including how they deal with the prescientific.

5.1 Direct Frequentism

Most statisticians are frequentist.[39] They test statistical hypotheses. Their main mathematical outcomes are determinations about statistical significance, p-values, confidence intervals and so-called "empirical estimates."

Here, I will focus on frequentist testing, which is a probabilistic adaptation of logical disproof by contradiction.[40] Logically, one may disprove hypothesis H by observing an event E that could not occur if H were true.[41] One may regard frequentist testing as a weaker form of that logic: if data x are obtained which would be unlikely if H were true, then H is "rejected." Once a test is designed, its results depend only on the data, so they are often called "objective."

In this "objectivistic" frequentist tradition, one scientific conclusion is prescribed if the tested H is rejected, and another is prescribed otherwise.[42] Neyman and Pearson proposed this approach "as an effort to introduce some hard logic, as opposed to informal judgment, to ideas of significance testing." [43] They reasoned that, if an event *E* is obtained that would be rare if *H* were true, then one may act as if *H* is false.

I will call this tradition "direct frequentism" because its decisions proceed based purely on the test results. It leaves no room for prescientific considerations, such as experts' judgments and the outcomes of previous studies. It makes statistics an "automatic technology," [44] rather than the "artful science" envisioned by R.A. Fisher.[45]

5.2 Indirect Frequentism

A popular "subjectivistic" alternative to direct frequentism might be called "indirect frequentism."[46] Sterne and Smith,[47] for instance, would have us

> move from the idea that results are significant or non-significant to the interpretation of findings in the context of the type of study and other available evidence. . . Here the prior evidence is simply the results of previous studies of the same issue. Other forms of evidence are, of course, admissible: findings from domains as different as animal studies and tissue cultures on the one hand and secular trends and ecological differences in human disease rates on the other will all influence a final decision as to how to act in the light of study findings.

That is, evaluation of any hypothesis *H* requires a thorough evaluation of all available evidence about *H*, of which statistical test results, though necessary, are only elements.

Mclean[48] promotes a similar stance. First, he criticizes what I have called the direct frequentist practice of considering results

> to have been established, or not, according to whether they [are] significant at the 5% level, or at the 1% level. This misuse of hypothesis testing has led to results with p=0.049 accepted, and those with p=0.051 rejected, and is one of the major reasons for the controversy of the last half century.

He then pinpoints the essence of what I mean by indirect frequentism: "Statistics is about judgment, and ironclad rules such as these discount the role of judgment."

In indirect frequentism, any given statistical test outcome does not necessarily lead to a particular conclusion about the tested hypothesis.[49] The outcome is rather combined informally with prescientific considerations, including estimates of magnitudes of effects,[50] results of other studies[51] and the "prior plausibility of the hypothesis under test and the desired impact of the results."[52]

Indirect frequentism's judgments differ from direct frequentism's "automatic" process of evaluating hypotheses based ostensibly on mathematical and logical rules. Whereas direct frequentism interprets test results rigidly, indirect frequentism interprets them freely and creatively.

Indirect frequentists usually construe test results as measures of "evidence" against the tested hypothesis *H*.[53] Burdette and Gehan[54] provide some typical guidance for this by suggesting cut-points, albeit arbitrary, for characterizing the p-value as "very strong," "moderate," etc. evidence. Once a researcher has so classified the p-value, he or she often has a surer feeling, justified or not, for how to combine it with prescientific considerations, using the contextual "judgment" Sterne, Smith and Mclean believe statistics is about.

Direct and indirect frequentism have co-existed for some time, and may enjoy roughly the same popularity presently.[55]

5.3 Frequentist Test Results are Deductive

Frequentist results are deductive statements about data x given hypotheses H, rather than inductive statements about hypotheses. For instance, the "95%" of the term "95% confidence interval for the mean" refers to a pre-experimental probability that the interval eventually will contain the mean. It is not a post-experimental probability about the mean itself.[56]

Therefore, deriving inductive meaning from them usually involves misinterpretation.[57] For example, most conclusions based on frequentist results comprise some form of the "transposition of conditioning,"[58] which has general structure[59] $Pr(H|x)=Pr(x|H)$. We commit that transposition if we read $Pr(+T|H_1)$ as $Pr(H_1|+T)$ in the example above, or if we read the p-value as $Pr(H|x)$.[60] Another common manifestation of it is maximum likelihood estimation. This estimation assumes that, given x, the θ maximizing the deductive $f(x|\theta)$ is itself inductively most credible, as if it maximized $\pi(\theta|x)$.[61]

Due to the disparity between frequentist statements and scientific induction, frequentist testing has been mistrusted, misappropriated and criticized throughout its history. Thompson[62] cataloged over 400 references against its "indiscriminate uses." Its results are misunderstood as often as understood.[63] Oakes' studies[64] concluded that even experienced scientists, and not only students, usually misread them.

Even the earliest promoters of frequentist testing dissociated it from scientific induction. Neyman and Pearson[65] warned "[n]o test based upon a theory of probability can by itself provide any valuable evidence of the truth or falsehood of a hypothesis." Similarly, Fisher[66] cautioned against measuring inductive evidential strength using deductive measures of rareness such as error rates and p-values: "The infrequency with which, in particular circumstances, decisive evidence is obtained, should not be confused with the force, or cogency, of such evidence."

Statistical testing is insufficient for the induction to which statistics aspires. It is widely practiced not due to any usefulness or correctness, but, as I will show, due to its consonance with scientism and subjectivism.

5.4 Objective and Subjective Bayesianism

Bayesian statisticians take Bayes' theorem, a form of which is

$$\pi(\theta \mid x) = \frac{f(x \mid \theta)\,\pi(\theta)}{f(x)},$$

"as a guide for revising belief in the light of new experimental evidence,"[67] where

> $\pi(\theta|x)$ = the "posterior," or post-experimental probability distribution function of the unknown parameter θ given data x,
> $f(x|\theta)$ = the likelihood of x given θ,
> $\pi(\theta)$ = the "prior," or pre-experimental probability distribution function of θ, and
> $f(x)$ = the unconditional probability distribution function of x.

The main inferential bayesian results are probabilities of hypotheses, and are calculated as follows. Consider any interval $[a,b]$ of possible values for θ. The prior (posterior) belief that $\theta \in [a,b]$ is represented by the area under the prior (posterior) distribution between a and b.

Probabilities of hypotheses are straightforward, inductive statements about unknown statistical parameters, fitting well our definition of statistical inference. Yet, only a minority of statisticians identify themselves as bayesians[68] and, as recently as 1999, most statistics departments did not even regularly offer bayesian courses.[69]

Traditionally, bayesians are classified as objective or subjective, according their methods of developing priors. Objective bayesians form priors utilizing mathematical principles, such as uniformity, "maximum entropy" or "non-informativity." These principles do increase inter-subjective agreement on priors. Such agreement is not complete, though, because different mathematical principles often lead to conflicting priors.

Subjective bayesians, on the other hand, generate priors making use of whatever prescientific background information that is available about parameters. All the considerations mentioned in "Indirect Frequentism" above about expected magnitudes of effects, conclusions of previous studies and so forth, may play a role. Not everyone agrees, though, about the admissible types of information or the ways they should be combined. For example, scientific empiricists form priors based primarily on experimentation, whereas scientific rationalists emphasize theorizing and introspection.[70]

6. Statistical Approaches on the Prescientific

One can discern various attitudes about the prescientific in the above paradigms. They vary along two "dimensions," at least:

6.1 Usage versus Disqualification of Prescientific Information

Indirect frequentism and subjective bayesianism openly incorporate the prescientific in conclusions. Direct frequentism and objective bayesianism, in contrast, rule it out. Plainly, these preferences can influence scientific conclusions based on statistical outcomes. If, for example, a prescientific background and experimental data conflict, then outcomes based on subjective priors can differ from those based on objective priors.

Dunnett[71] illustrates these potential differences clearly. He offers advice on interpreting test results that seem to say a particular high drug dose is less efficacious than a small dose. He claims that any inference regarding "whether a real treatment effect has been demonstrated, which for some reason is not manifested at the higher dose level, would depend on the experimenter's prior knowledge regarding the properties of this particular drug." Plainly, "indirectly" regarding test results merely as evidence to be combined with pharmacological knowledge, and "directly" regarding them as decisive, can lead to diverging scientific conclusions.

Precisely because of direct frequentism's disqualification of the prescientific, many prefer this paradigm to bayesianism, despite its silence with respect to induction. Its results, though obscure and usually misunderstood, are "objective,"[72] meaning that they can be calculated based on empirical data alone. For, to non-bayesians, the most objectionable feature of bayesianism "is the use of the notion of a 'prior' probability. This concept has been severely criticized . . . as representing subjective bias."[73] Hampel,[74] for instance, acknowledges that "[t]he Bayesian paradigm would be ideal, were it not for the choice of the first prior distribution." Frequentists take exception, then, to bayesians' explicit use of extra-statistical information that, they believe, detracts from objectivity.[75]

6.2 Intra-analytic versus Post-analytic Use of Prescientific Information

Scientific conclusions based on statistical results can vary depending also on whether prescientific information is incorporated during analysis or after analysis.

Subjective bayesianism forms priors based on this information. It then mathematically combines these priors with quantitative data, intra-analytically incorporating the prescientific in the statistical results. The resulting posteriors are considered final statements of belief. They function normatively, indicating what people should believe given the priors and the data. Quantitative and logical laws are given the "last word" in conclusions.

In indirect frequentism, though, the prescientific does not affect statistical outcomes. It is rather brought together post-analytically with those outcomes, to affect only final scientific conclusions. The "last word" is handed over to intuition and judgment.

Attitudes about the role of prescientific information in statistics, then, influence theorizing in the sciences in which statistics is applied. This motivates us to consider what a biblically consistent model for this role might be.

7. Statistical Paradigms and the Prescientific

7.1 Direct Frequentism, Objective Bayesianism and Scientism

Now, we have seen that direct frequentism and objective bayesianism calculate their conclusions independently of the prescientific, and offer them as objective quantitative-logical determinations of what is fiducially credible. They regard these abstract statistical results as the "genuine"[76] guides for scientific beliefs, rationalizing a reduction of fiduciary functioning to logical and quantitative functioning. Therefore, these statistical paradigms violate the PLI's "irreducibility" principle. They utilize statistical outcomes not to build on or add to prescientific experience, but to "replace" that experience.[77] Therefore, they seem consistent with the scientism mentioned above.

This consistency accounts for the popularity of direct frequentism that, I have shown, does not meet the inductive needs of statistics.

Hampel[78] illustrates this stance, dogmatically banishing prescientific experience from statistics. He claims "inference based on a start with total ignorance is trustworthy even if wasting a bit of information, while inference based on a bayesian prior is shaky and unreliable by an unknown amount." Any claim, however, that "total ignorance" is categorically more trustworthy than prescientific experience cannot arise from an appreciation of multi-aspectual reality as it is.[79] It reflects rather a commitment to the autonomy of statistics from prescientific presuppositions. If the prescientific is that "shaky and unreliable," then civil engineers should be retired after a year or so, so their accumulated knowledge won't jeopardize diagnoses of bridge problems!

7.2 Indirect Frequentism and Subjectivism

Indirect frequentist is strikingly consistent with the particular subjectivism portrayed above. Its analyses begin essentially as do direct frequentist and objective bayesian ones: by generating abstract, inter-subjectively verifiable test results. Multi-aspectual prescientific experiences are assumed once more to "lack 'general validity' or 'objectivity.'"[80] This frequentism regards statistical test results as "un-problematical data,"[81] and ignores the problems with them mentioned above. In these respects, indirect frequentism resembles both scientism as well as the subjectivism we have noted.

Once test results are generated, indirect frequentism takes on a character quite distinct from scientism, though very much in agreement with the subjectivism we have noted. It leaves individual scientists to "order [test results] into an intelligible experience" using independent, informal reasoning. They decide for themselves what the entire body of evidence signifies.

In this manner, indirect frequentism attempts to preserve the freedom of the researcher. It avoids what the above subjectivism sees as "transgressing into the normative sphere,"[82] by not stipulating what is appropriate to believe based on statistical results. Indirect frequentism is fashionable, therefore, due to its subjectivist features, and despite (or perhaps because of) its failure to substantiate inductive statements about hypotheses.

7.3 Subjective Bayesianism and Enhancing Subject-Object Relations

The PLI suggests that statistical analysis must make prescientific experience the foundation for inference, in a way that allows scientific "objects" to place "limitations" on "the subject who does the knowing."[83]

Subjective bayesianism seems to meet these specifications. First, its priors constitute a device for building on the prescientific. They allow one to embed inference into multi-aspectual reality, before the latter is "dissected." For instance, a physicist may be aware of the physical properties of a particular set of things. This awareness can affect his/her quantitative, logical and fiduciary degree of belief that they will behave a certain way. That belief can be summarized in a subjective prior.

Second, subjective bayesianism's prior and posterior probabilities are consistent with the PLI in that they are direct (inductive) statements of certainties about scientific objects. They identify mathematical "limitations" on subjects' quantitative-logical-fiduciary functions.

Subjective bayesian inference is, in summary, a mathematical process of changing the fiduciary subject-object relations between scientists and their subject matters. It assumes the scientist possesses prescientific degrees of belief. It modifies these degrees from their pre-experimental to their post-experimental states, using quantitative evidence. It does not try to create new relations with scientific objects but, consistent with the PLI, enhances relations that exist with them already.

8. Conclusions

The PLI's treatment of the prescientific entails three principles, at least, about the subjective and objective bounds of statistics:

1) Multi-aspectual, prescientific experience, rather than abstraction, is the proper foundation for inference.
2) Contrary to scientism, we cannot expect empirical, quantitative data to create knowledge "ex nihilo," out of nothing. We must instead combine these data with prescientific knowledge.
3) Yet, contrary to subjectivism, inference must identify relevant objective limitations on belief. This is possible if we respect the mutual irreducibility and coherence between the aspects of reality.

In these respects, subjective bayesianism seems more compatible with the PLI than do other statistical paradigms. To the extent the PLI is accepted as a Christian philosophy, "making every [statistical] thought obedient to Christ" requires serious consideration of this statistical paradigm.

The PLI offers, in addition to these principles, an explanation for the ubiquitous custom of reading inductive meaning into deductive probabilities: this custom is consistent with the ideals of scientism and subjectivism in science.

Despite these conclusions, we cannot pretend to have expounded a "magic formula" for conducting statistics Christianly. Indeed, we'll never be able to define, once and for all, a Christian attitude towards any particular science: "Scientific activity belongs to the task given to man at the creation, and is a task that will forever remain incomplete, tentative, developing."[84] What we can attempt, though, is to contribute to the overall historical development of statistics, in a manner to some extent consistent with God's design.[85]

I have hoped, then, not only to answer questions in this essay, but to raise them. In particular, we need to investigate what the PLI may indicate about the following topics:

What are degrees of belief? What are their properties?
What are the mechanisms for measuring uncertainty, viz., "translating" the uncertain knowledge of the other sciences into subjective bayesian priors?
Can we hope to quantify all forms of uncertainty? In other words, is all uncertainty "quantitatively relevant?"[86]

A recurring biblical theme is that, while God's law constrains us, it also frees us to function as we were designed to function (Pr 3:18). I take this to mean that, as we work out the above and other statistical topics, we can expect to discern the liberating effects of searching out God's design in all our activities.[87]

Endnotes

[1]David Huntsberger, *Elements of Statistical Inference.* (Boston: Allyn and Bacon, 1961), 2.
[2]Robert Peto, et al., "Design and analysis of randomized clinical trials requiring prolonged observation of each patient (Part I)," *Br. J. Cancer 34* (1976), 585-612, 595.
[3]Jan Geertsema, "A Christian view of the foundations of statistics," *Perspectives on Science and Christian Faith 39:3* (1987): 158-164.
[4]E.g., in Kalsbeek, *Contours of a Christian Philosophy*, (Toronto: Wedge, 1975).
[5]By *applied* statistical inference, I refer to inference as it is applied in various other sciences (physics, biology, sociology. . .). This paper is not primarily to show that subjectivity is unavoidable in statistical inference (although that is true). It shows rather the types of objectivity and subjectivity that are both possible and necessary in statistics.
[6]Bernard Zylstra, "Introduction." In Kalsbeek, 14-33, 31.
[7]Roy Clouser gives a fuller account of the PLI in his *Myth of Religious Neutrality*, reprinted 2001 (Notre Dame, Indiana: Notre Dame Press, 1991).
[8]In the PLI, an *object* is, among other meanings, ". . . a thing; anything known by a knower. . ." A *subject* "can refer to . . . a person, a knower. . ." [Hendrik Hart, *Understanding our World: An Integral Ontology* (Lanham, MD: University Press of America, 1984), 222].
[9]Herman Dooyeweerd, *Roots of Western Culture* (Toronto: Wedge, 1979), 33.
[10]Herman Dooyeweerd, *A New Critique of Theoretical Thought*, translated by DH Freeman and H DeJongste (Philadelphia: Presbyterian and Reformed Publishing Company, 1955), Vol II, 103.
[11]Dooyeweerd, *New Critique*, Vol II, 82.
[12]Clouser, 217.
[13]Ibid.
[14]H. Evan Runner, *The Relation of the Bible to Learning* (Jordan Station, Ontario: Paideia Press, 1982), 151.
[15]Herman Dooyeweerd, *Encyclopedia of the Science of Law*, translated by R.D. Knudsen (New York: Mellen, 2002), Vol 1, 22-23.
[16]Martin Rice, "What is a Science?" In Strauss and Botting (eds), *Contemporary Reflections on the Philosophy of Herman Dooyeweerd* (New York: Mellen, 2000).
[17]D.F.M. Strauss, *Paradigms in Mathematics, Physics and Biology: their Philosophical Roots* (Bloemfontein: Tekskor, 2001), 13.
[18]Clouser, 58.
[19]There is a view that statistics is a strictly mathematical science, which explains why statistics is often taught in universities' mathematics departments. The conception of statistics I present here conflicts with that view, though. Mathematics focuses on the quantitative aspect, but statistics focuses on the quantitative, logical and fiduciary ones.
[20]Runner, 125.
[21]Runner, 125.
[22]Dooyeweerd, *Encyclopedia*, 39-40.
[23]Dooyeweerd, *Encyclopedia*, 42.
[24]Clouser, 208.
[25]Dooyeweerd, *Encyclopedia*, 42.
[26]Hart, 232.
[27]Clouser, 214.
[28]Precise definitions of applied statistics are surprisingly rare; even most introductory statistical texts don't describe statistics globally, but jump right into prescribing methods for doing it. Others state their

"definitions" in terms of examples of what statisticians are seen doing in the workplace. In any case, my definition is consistent with many such depictions.

[29]James Berger, *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. (New York: Springer, 1985), 7.

[30]Ronald Fisher, *Statistical Methods for Research Workers*, 1st ed. (Edinburgh: Oliver and Boyd, 1925), 1.

[31]Geertsema.

[32]Huntsberger, 5.

[33]Robert Anderson, *Practical Statistics for Analytical Chemists*, 1st ed. (New York: Van Nostrand Reinhold, 1987), 3.

[34]As is often concluded in diagnostic medicine, according to Frank Davidoff, "Standing statistics right side up," *Annals of Internal Medicine 130*:12 (1999): 1019-1021, 1019.

[35]Here, □ is a variable, not the number 3.14159. . .

[36]By Kolmogorov's definition of conditional probability,

$$Pr(H_1|+T)=Pr(+T, H_1) / Pr(+T) \text{ and}$$
$$Pr(+T|H_1)= Pr(+T, H_1) / Pr(H_1),$$

so that

$$Pr(+T, H_1)= Pr(+T|H_1) Pr(H_1).$$

Thus,

$$Pr(H_1|+T)=Pr(+T| H_1)Pr(H_1) / Pr(+T).$$

Now,

$$Pr(+T)=Pr(+T|H_1)Pr(H_1) + Pr(+T|H_0)Pr(H_0)$$
$$=0.80 \pi + 0.05(1- \pi) = 0.05+0.75\pi.$$

Therefore,

$$Pr(H_1|+T)=0.8\pi / (0.05+0.75\pi).$$

[37]Admittedly, valid reasons for deviating from $Pr(H_1|+T)$ could exist. One might, for example, possess a pre-test probability about H_1 that differs from □, or suspect that the collection of the datum $+T$ was flawed. Nonetheless, if one's pre-test probability is indeed □ and one trusts the testing procedures, then one's certainty about H_1 is this $Pr(H_1|+T)$, in a *normative* sense. This type of meaning is not conveyed by frequentist results.

[38]Davidoff, 1019.

[39]J Martin Bland and Douglas Altman, "Bayesians and frequentists." *British Medical Journal 317* (1998): 1151-1160; Robert Matthews, *Facts versus factions: The use and abuse of subjectivity in scientific research.* (Cambridge, UK: European Science and Environment Forum, 1998).

[40]Carmen Batanero, "Controversies around the role of statistical tests in experimental research," *Mathematical Thinking and Learning 2*:1-2 (2003): 75-98.

[41]Karl Popper, *The Logic of Scientific Discovery*. (London: Rutledge, 1959).

[42]Huntsberger, 141-147; Anderson, 52; Y Leon Maksoudian, *Probability and Statistics with Applications* (Scranton, Pennsylvania: International Textbook Company, 1969), 221-232; Robert Hogg and Elliot Tanis, *Probability and Statistical Inference*, 4th ed. (New York: Macmillan, 1993), 394-399.

[43]Oscar Kempthorne, "Of what use are tests of significance and tests of hypotheses?" *Communications in Statistics - Theory and Methods A5*: 8 (1976): 763-777, 764.

[44]Egbert Schuurman, "Beyond the empirical turn: responsible technology," Egbert Schuurman, http://home.wxs.nl/sch_art.htm (accessed Dec 5, 2003); Jonathan Sterne, "Teaching hypothesis tests - time for significant change?" *Statistics in Medicine 21*(2002): 985-994, 991.

[45]Ronald Fisher, *Statistical Methods for Research Workers*, 9th ed. (Edinburgh: Oliver and Boyd, 1944).

[46]Direct and indirect frequentism could also be called "naive" and "sophisticated" frequentism, for they correspond, in somewhat striking ways, to the "naive falsificationism" and "sophisticated falsificationism" philosophies of science promoted by Karl Popper and his student Imre Lakatos [Imre Lakatos, *Criticism and the Growth of Knowledge,* (New York: Cambridge University Press, 1970), 91-195; from Theodore Schick, Jr., ed., *Readings in the Philosophy of Science*, (Mountain View, CA: Mayfield Publishing Co: 2000)]. In fact, Colin Howson (personal communication) has suggested that Popper found some of his inspiration for falsificationism in the frequentist "hero of 20th Century statistics," RA Fisher.

[47]Jonathan Sterne and George Smith, "Sifting the evidence - What's wrong with significance tests?" *British Medical Journal 322* (2001):226-231.

[48]Alan McLean, “On the nature and role of hypothesis tests,” Working Paper 4/2001 (Australia: Monash University, Department of Econometrics and Business Statistics), 15.
[49]John Ware, Frederick Mosteller and Joseph Ingelfinger, “P-values.” In John Bailar and Frederick Mosteller (eds), *Medical Uses of Statistics* (Waltham, Massachusetts: NEJM Books, 1986), 158.
[50]Frank Yates, “The influence of ‘Statistical Methods for Research Workers’ on the development of statistics,” *Journal of the American Statistical Association 46* (1951): 19-34, 32; Institute for Clinical Harmonization (ICH), “ICH Guideline E9: Statistical Principles for Clinical Trials,” ICH, http://www.chinagmp.net/ICH/ich-files/E9.PDF (accessed 24 Aug 2003), paragraph 2.1.2; Steven Goodman, “Toward evidence-based medical statistics. 1: The p-value fallacy,” Annals of Internal Medicine 130 (1999): 995-1004, 998; European Agency for the Evaluation of Medicinal Products (EMEA), “Points to Consider on Adjustment for Baseline Covariates,” EMEA, http://www.emea.eu.int (accessed 27 Apr 2003), 4; Jacob Cohen, “Things I have learned (so far),” *Journal of the American Psychological Association 45*: 12 (1990): 1304-1312.
[51]ICH 1998 E9, Section 1.2.
[52]ICH 1998 E9 Sections 3.5 and 2.1.2; Sterne and Smith.
[53]Batanero, 78; Sterne, 990; Hilary Term, “Descriptive Statistics for Research,” Institute for the Advancement of University Learning & Department of Statistics (2002), 14; Goodman, “P-value fallacy,” 998; Gudmund Iversen and Mary Gergen, *Statistics: The Conceptual Approach* (New York: Springer, 1997), 271.
[54]Walter Burdette and Edmund Gehan, *Planning and Analysis of Clinical Studies* (Springfield, IL: Charles C. Thomas, 1970), 9.
[55]Sterne, 990, provides evidence that both direct and indirect frequentism constitute popular positions at present, at least in medical schools in the U.K. Sixteen such schools were asked to agree or disagree, on a questionnaire, with two direct frequentist statements (among other statements) about the p-value p. Fully half (8) of the schools agreed with the first, and 7 agreed with the second, pointing to a nearly even split between direct and indirect frequentism. Cohen (1990) actually alludes to a dominance of direct frequentist attitudes. After commending to us an indirect frequentist stance, he laments that a majority of scientists interpret test results directly:

> An essential ingredient in the research process is the judgment of the scientist. He or she must decide by how much a theoretical proposition has been advanced by the data, just as he or she decided what to study, what data to get, and how to get it. I believe that statistical inference applied with informed judgment is a useful tool in this process. . . The implications of the things I have learned (so far) are not consonant with much of what I see about me as standard statistical practice. The prevailing yes-no decision at the magic 0.05 level from a single research is a far cry from the use of informed judgment. Science simply doesn't work that way.

The important point here is that direct and indirect frequentism have co-existed for some time and, while the indirect sort may be exerting more influence at present than, say, 50 years ago, the direct sort is by no means dead.
[56]David Salsburg, *The Lady Tasting Tea*, (New York: Henry Holt, 2002), 117-124.
[57]Both direct and indirect frequentist statements are unjustified, as I show, yet, if one does not interpret test results directly or indirectly, then what can one say about the tested hypothesis given the data?
[58]Berger, *Statistical Decision Theory*, 120; James Berger and Thomas Selke, “Testing a Point Null Hypothesis: The irreconcilability of P-values and evidence,” *Journal of the American Statistical Association 82*: 397 (1987): 112-122; Davidoff; Goodman, “P-value fallacy;” Steven Goodman, “Toward evidence-based medical statistics. 2: The Bayes Factor,” *Annals of Internal Medicine 130* (1999), 1005-1013; Cohen; Matthews; Batanero, 12.
[59]By Kolmogorov’s definition of conditional probability,

$$Pr(H|x)=Pr(x|H)[Pr(H) / Pr(x)],$$

so that this transposition of conditioning $Pr(H|x)=Pr(x|H)$ is inaccurate quantitatively to the extent that the quotient in brackets differs from unity.
[60]Goodman, “P-value fallacy,” 997.
[61]It is true that a well-taught introductory statistics course will deal with the danger of the transposition of conditioning. However, “[s]tudents notoriously find hypothesis testing a difficult subject to learn properly. Although it is not difficult for students to carry out a classroom (or examination) hypothesis test exercise, it is difficult to convince a teacher that the students really understand what they have learnt. There is

considerable research into what students understand by, for example, 'level of significance' which demonstrates considerable confusion. . . .Given the confusion in the research community, this is not surprising" (McLean, 2001). That is to say, despite all efforts, most frequentist results are interpreted using the transposition and its variants [Daniel Wright, "Making friends with your data: Improving how statistics are conducted and reported." *British Journal of Educational Psychology 73* (2003), 123-136, 124].

[62]William Thompson, "402 Citations Questioning the Indiscriminate Use of Null Hypothesis Significance Tests in Observational Studies," William Thompson, http://biology.uark.edu/Coop/thompson5.html (accessed April 18, 2003).

[63]University of Stirling. "Oakes' Study," University of Stirling, http://www.stir.ac.uk/departments/humansciences/psychology/mscp&h/resm/lecture2/tsld003.htm (accessed April 24, 2003); Mclean; Berger, *Statistical Decision Theory*, 123; George Diamond and James Forrester, "Clinical trials and statistical verdicts: Probable grounds for appeal," *Annals of Internal Medicine 98* (1983): 385-394, 385.

[64]Michael Oakes, *Statistical Inference* (Chichester: Wiley, 1986).

[65]Goodman, "P-value fallacy," 998.

[66]Ronald Fisher, *Statistical Methods and Scientific Inference*, 2nd edition (New York: Hafner, 1959), 82.

[67]Kenneth Schaffner, "Ethically optimizing clinical trials." In *Bayesian Methods and Ethics in a Clinical Trial Design*, edited by J.B. Kadane (New York: Wiley, 1996), 19-63.

[68]Sterne, 990.

[69]James Berger, "Bayesian analysis: A look at today and thoughts of tomorrow." Prepared as a vignette for *Journal of the American Statistical Association*, James Berger, http://www.isds.duke.edu/~berger/papers/99-30.html (accessed February 16, 2004), Section 1.

[70]Geertsema.

[71]Charles Dunnett, "New tables for multiple comparisons with control," *Biometrics 20* (1964): 482-491, 485.

[72]Matthews.

[73]Schaffner, 42.

[74]F. Hampel, "On the foundations of statistics: A frequentist approach," Research Report No. 85, CH-8092 (Zurich, Switzerland: Seminar fur Statistik, 1998), Section 1.3.

[75]Geertsema; Robert Matthews, "Methods for assessing the credibility of clinical trial outcomes," *Drug Information Journal 35* (2001): 1469-1478, 1473.

[76]Runner, 125.

[77]Runner, 125.

[78]Hampel, Section 2.6.

[79]Indeed, "total ignorance" would stop frequentism dead in its tracks. For, frequentist inference depends on assumptions regarding what its deductive statements about data signify inductively about hypotheses. Furthermore, such ignorance is impossible to express in an objective bayesian prior distribution. Even a uniform prior claims that the probability that □□*I* is the same for any interval *I* of fixed length. This is hardly a statement of total ignorance.

[80]Dooyeweerd, *Encyclopedia*, 39.

[81]Dooyeweerd, *Encyclopedia*, 39.

[82]Dooyeweerd, *Encyclopedia*, 42.

[83]Hart, 232; Clouser, 214.

[84]Wytze Brouwer, "Christian Commitment and Scientific Theories" (Toronto: Association for the Advancement of Christian Scholarship, no date), 1.

[85]Dooyeweerd, *New Critique*, Vol I, 118-119.

[86]Russ Wolfinger, personal communication.

[87]The author thanks Bruce Wearne, Roy Clouser and Russ Wolfinger for their collaboration and for reviewing some manuscripts of this submission.

Bibliography

Anderson, R. 1987. *Practical Statistics for Analytical Chemists*, 1st ed. New York: Van Nostrand Reinhold.

Batanero, C. 2003. Controversies around the role of statistical tests in experimental research. *Mathematical Thinking and Learning 2*:1-2: 75-98.

Berger, J. 1985. *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. New York: Springer.

Berger, J. and Selke, T. 1987. Testing a Point Null Hypothesis: The irreconcilability of P-values and evidence. *Journal of the American Statistical Association 82*: 397: 112-122.

Berger, J. 1999. Bayesian analysis: A look at today and thoughts of tomorrow. Prepared as a vignette for *Journal of the American Statistical Association*, James Berger, http://www.isds.duke.edu/~berger/papers/99-30.html (accessed February 16, 2004), Section 1.

Bland, J. and Altman, D. 1998. Bayesians and frequentists. *British Medical Journal 317*: 1151-1160.

Brouwer, W. no date. Christian Commitment and Scientific Theories. Toronto: Association for the Advancement of Christian Scholarship.

Burdette, W. and Gehan, E. 1970. *Planning and Analysis of Clinical Studies.* Springfield, IL: Charles C. Thomas.

Clouser, R. 1991. *Myth of Religious Neutrality*, reprinted 2001. Notre Dame, Indiana: Notre Dame Press*.*

Cohen, J. 1990. Things I have learned (so far). *Journal of the American Psychological Association 45*: 12: 1304-1312.

Davidoff, F. 1999. Standing statistics right side up. *Annals of Internal Medicine 130*:12: 1019-1021.

Diamond, G. and Forrester, J. 1983. Clinical trials and statistical verdicts: Probable grounds for appeal. *Annals of Internal Medicine 98:* 385-394.

Dooyeweerd, H. 1955. *A New Critique of Theoretical Thought*, translated by DH Freeman and H DeJongste. Philadelphia: Presbyterian and Reformed Publishing Company.

Dooyeweerd, H. 1979. *Roots of Western Culture.* Toronto: Wedge.

Dooyeweerd, H. 2002. *Encyclopedia of the Science of Law*, translated by R.D. Knudsen. New York: Mellen.

European Agency for the Evaluation of Medicinal Products (EMEA), Points to Consider on Adjustment for Baseline Covariates, EMEA, http://www.emea.eu.int (accessed 27 Apr 2003).

Dunnett, C. 1964. New tables for multiple comparisons with control. *Biometrics 20*: 482-491.

Fisher, R. 1925. *Statistical Methods for Research Workers.* 1st ed. Edinburgh: Oliver and Boyd.

Fisher, R. 1959. *Statistical Methods and Scientific Inference.* 2nd ed. New York: Hafner.

Fisher, R. 1944. *Statistical Methods for Research Workers.* 9th ed. Edinburgh: Oliver and Boyd.

Goodman, S. 1999. Toward evidence-based medical statistics. 1: The p-value fallacy. *Annals of Internal Medicine 130*: 995-1004.

Goodman, S. 1999. Toward evidence-based medical statistics. 2: The bayes factor. *Annals of Internal Medicine 130*, 1005-1013.

Geertsema, J. 1987. A Christian view of the foundations of statistics. *Perspectives on Science and Christian Faith 39:3*: 158-164.

Hampel, F. 1998. On the foundations of statistics: A frequentist approach. Research Report No. 85, CH-8092. Zurich, Switzerland: Seminar fur Statistik.

Hart, H. 1984. *Understanding our World: An Integral Ontology.* Lanham, MD: University Press of America.

Hogg, R. and Tanis, E. 1993. *Probability and Statistical Inference*, 4th ed. New York: Macmillan, 394-399.

Huntsberger, D. 1961. *Elements of Statistical Inference.* Boston: Allyn and Bacon.

Institute for Clinical Harmonization (ICH), ICH Guideline E9: Statistical Principles for Clinical Trials. ICH, http://www.chinagmp.net/ICH/ich-files/E9.PDF (accessed 24 Aug 2003).

Iversen, G. and Gergen, M. 1997. *Statistics: The Conceptual Approach.* New York: Springer.

Kalsbeek, L. 1975. *Contours of a Christian Philosophy*. Toronto: Wedge.

Kempthorne, O. 1976. Of what use are tests of significance and tests of hypotheses. *Communications in Statistics - Theory and Methods A5*: 8: 763-777.

Lakatos, I. 1970. *Criticism and the Growth of Knowledge*. New York: Cambridge University Press.

Maksoudian, Y. 1969. *Probability and Statistics with Applications.* Scranton, Pennsylvania: International Textbook Company.

Matthews, R. 1998. *Facts versus factions: The use and abuse of subjectivity in scientific research.* Cambridge, UK: European Science and Environment Forum.

Matthews, R. 2001. Methods for assessing the credibility of clinical trial outcomes. *Drug Information Journal 35*: 1469-1478.

McLean, A. 2001. On the nature and role of hypothesis tests. Working Paper 4/2001 (Australia: Monash University, Department of Econometrics and Business Statistics).

Oakes, M. 1986. *Statistical Inference.* Chichester: Wiley.

Popper, K. 1959. *The Logic of Scientific Discovery*. London: Rutledge.

Peto, R. et al. 1976. Design and analysis of randomized clinical trials requiring prolonged observation of each patient (Part I). *British Journal of Cancer 34*, 585-612.

Rice, M. 2000. What is a Science. In Strauss and Botting (eds), *Contemporary Reflections on the Philosophy of Herman Dooyeweerd.* New York: Mellen.

Runner, H. 1982. *The Relation of the Bible to Learning.* Jordan Station, Ontario: Paideia Press.

Salsburg, D. 2002. *The Lady Tasting Tea*. New York: Henry Holt.

Schaffner, K. 1996. Ethically optimizing clinical trials. In *Bayesian Methods and Ethics in a Clinical Trial Design*, edited by J.B. Kadane. New York: Wiley.

Schick, Jr., T., ed., 2000. *Readings in the Philosophy of Science.* Mountain View, CA: Mayfield Publishing Co.

Schuurman, E. 2003. Beyond the empirical turn: responsible technology. Egbert Schuurman, http://home.wxs.nl/sch_art.htm (accessed Dec 5, 2003).

Sterne, J. 2002. Teaching hypothesis tests - time for significant change. *Statistics in Medicine 21*: 985-994.

Sterne, J. and Smith, G. 2001. Sifting the evidence - What's wrong with significance tests. *British Medical Journal 322*:226-231.

Strauss, D.F.M. 2001. *Paradigms in Mathematics, Physics and Biology: their Philosophical Roots.* Bloemfontein: Tekskor.

Term, H. 2002. Descriptive Statistics for Research. Institute for the Advancement of University Learning & Department of Statistics.

Thompson, W. 2003. 402 Citations Questioning the Indiscriminate Use of Null Hypothesis Significance Tests in Observational Studies. William Thompson, http://biology.uark.edu/Coop/thompson5.html (accessed April 18, 2003).

University of Stirling. 2003. Oakes' Study. University of Stirling, http://www.stir.ac.uk/departments/humansciences/psychology/mscp&h/resm/lecture2/tsld003.htm (accessed April 24, 2003).

Ware, J., Mosteller, F. and Ingelfinger, J. 1986. P-values. In John Bailar and Frederick Mosteller (eds), *Medical Uses of Statistics*. Waltham, Massachusetts: NEJM Books.

Wright, D. 2003. Making friends with your data: Improving how statistics are conducted and reported. *British Journal of Educational Psychology 73*, 123-136.

Yates, F. 1951. The influence of 'Statistical Methods for Research Workers' on the development of statistics. *Journal of the American Statistical Association 46*: 19-34.